Domain Parking: Not as Malicious as Expected

Leigh Metcalf, Jonathan Spring
netsa-contact@cert.org
CERT® Coordination Center, Software Engineering Institute
Publication CERTCC-2014-57

August 2014

Executive Summary

Domain parking is the practice of assigning a nonsense address to a domain when it is not in use in order to keep it ready for "live" use. This practice is peculiar because it indicates someone has administrative control over the domain name, does not have hardware ready to respond to requests, but wants the domain to appear active. A more appropriate response would seem to us to be that the domain does not exist. This mismatch between expected benign behavior (no such domain) and actual observed behavior (parking) made us suspicious. In this paper we discuss scalable detection methods for domain names parking on reserved IP address space, and then using this data set evaluate whether this behavior appears to be indicative of malicious behavior.

We find that during the month of January 2014 only 21,328 unique domains exhibited parking on reserved address space, out of over 610 million total unique observed domains. Thus, parking appears to be an uncommon Internet behavior with only 0.0035% of domains exhibiting parking on reserved IP addresses. Of these 21,328 domains, relatively few were observed listed on any of 16 domain black lists any time from January 1 to February 28, 2014. Only 1,563, or 7.3%, were listed in this time period. Therefore, we conclude that parking is a poor indicator of malicious activity, or at least not an indicator of any kind of malicious activity usually examined by any public list of malicious domain behavior.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 AUG 2014	**N/A**	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Domain Parking: Not as Malicious as Expected	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Metcalf /Jonathan Spring Leigh	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Software Engineering Institute Carnegie Mellon University Pittsburgh, PA 15213	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**10**	

1 Introduction

When a domain is "parked" on an IP address, the IP address to which the domain resolves is inactive or otherwise not owned by the domain owner. This is a common practice when a user first registers a domain – the registrar does not know what IP to supply as an answer, but supplying some answer prevents errors.

The domain name system permits a variety of different mechanisms which help provide resiliency to distributed architectures. Often these have legitimate uses, but malicious actors are equally able to adopt successful techniques. Usually the malicious use case is sufficiently different that the type of use can be teased apart. Suspicious domain name parking is no different; herein we present a method for finding it in historical passive DNS data.

Malicious actors seem to have adopted this technique for similar error suppression goals as the benign use case. Although it is suppression of different errors, such as evading detection before the number of infected machines reaches the desired number or while the command and control structure is not yet in place. We present a method for detection of domains that exhibit parking and a mechanism for distinguishing legitimate from suspicious use.

This parking destination, reserved IP space, is quite different from parking a domain on someone else's IP space. To our knowledge, there has been one study on parking domains for illicit ad revenue, which appears to happen on a large scale of 4 million domains [1]. However, from the authors' description this appears to be more like typosquatting (as described in Szurdi et. al. [2]) than resolution error suppression, as the authors describe the "dark side of domain parking" as monetized "whenever web users type in those domain names (probably accidentally) in the browser's address bar, the parking service resolves the domains to advertisement laden pages" [1, p. 1]. We are not aware of other studies of domain parking, except that some fast-flux identification algorithm studies cited domain parking as an obstacle [3, 4].

Parking on reserved IP space is sufficiently uncommon that it is somewhat difficult to find, at only 0.0035% of unique domains observed. This difficulty is not so much because it is infrequent but that the IP addresses commonly used for parking, such as the 127.0.0.0/8 block or those reserved in RFC 1918 [5] are also used for several other more common uses of the DNS, such as delivering real-time DNS black list results [6]. This introduces noise into any detection technique since it is not so simple as just finding domains that pointed to reserved address space at some time and then changed.

CIDR block	Justification
10.0.0.0/8	RFC 1918 [5]
127.0.0.0/8	RFC 1700 [9]
169.254.64.0/18	RFC 3927 [10]
172.16.0.0/12	RFC 1918 [5]
192.168.0.0/16	RFC 1918 [5]

Table 1: *Private IP address space*

2 Method

The main prerequisite for our method is a large source of passive DNS trace data. In order to calculate over large data volumes, we take several simplifying steps. Data is ingest in `nmsgtool` format [7], including source DNS server and precise time range the response was valid, at a rate of about 35 GB per day. Unique resource record sets (RRsets) are extracted from the DNS messages and extraneous fields are removed, leaving just the fields for `rname`, `TTL`, `type`, and `rdata` [8]. A list of unique RRsets per day based on these fields is approximately 2 GB in our data source.

Then, we load the RRsets with `type` of `A` (IPv4 answer) for January 2014 into a PostgreSQL database. The table has fields for the four RRset fields as well as day observed. Since RRsets are unique per day, if an identical RRset was observed on multiple days it will appear in the database for each day observed. This structure permits a course-grained time series view with enough data to detect patterns but enough summarization that calculation is practical.

In order to start our search for parking on private IP address space, we query the database for all RRsets where the `rdata` is in the IP set indicated in Table 1. Most of the results are not actually parking. Answers in private IP space are used to encode various kinds of non-location data, such as responses to lookups on DNSBLs, and for other administrative reasons in content distribution networks and hosting companies. We created a list by expert human analysis to remove these irrelevant domains from the results. Table 2 lists the second-level domains (SLDs) that were removed.

The process so far yields a list of RRsets with `rdata` in private IP space and `rname` domains that do not have a known use. We search for all other RRsets with the same domains in the `rname` field. Any results will have publicly routeable IP addresses, and thus at some point in the month have transitioned between private and routable IP address space. We consider these domains to have exhibited parking behavior on private IP address space.

abuseat.org	httpbl.org	schpider.com
ahbl.org	invaluement.com	senderscore.com
anubisnetworks.com	isipp.com	sonicwall.com
apews.org	ja.net	sophosxl.com
barracudacentral.org	jtripper.net	sorbs.net
bl.rptn.ca	junkemailfilter.com	spamcop.net
blocklist.de	kaspersky-labs.com	spameatingmonkey.net
bondedsender.org	lic.bizanga.net	spamhaus.net
borderware.com	lsu.edu	spamhaus.org
ciphertrust.net	mail-abuse.com	spamrats.com
clearswift.net	mailshell.net	spotilocal.com
cox.net	mailspike.net	srfsrs.com
dcrbl.com	mailspike.org	support-intelligence.net
ddnsbl.internetdefensesystems.com	manitu.net	surbl.org
device.trans.manage.esoft.com	mcafee.com	surfsrs.com
dns-rbl.at	microsoft.com	surriel.com
dnsbl.borderware.org	mooo.com	tornevall.org
dnsbl.inps.de	mozilla.org	trendmicro.com
dnsbl.it	msgsecurity.juniper.net	truncate.gbudb.net
dnsbl.justspam.org	nerd.dk	trustedsource.org
dnsresearch.us	nessus.org	uceprotect.net
dnswl.org	netvantasecurityportal.com	ucla.edu
drweb.com	njabl.org	ufl.edu
dsadns.net	nszones.com	uribl.com
dscwl.net	pacanka.com	validatorsearch.verisignlabs.com
dsintll.net	qualcomm.com	vircom.com
dsl.cantv.net	quorum.to	webcfs00.com
e5.sk	rating.cloudmark.com	webcfs01.com
enemieslist.com	rbl.esoft.com	webcfs02.com
eset.rs	rbl.zvelo.com	webcfs03.com
f1.dsmpd.net	sa.skype.net	wisc.edu
f1.dsusl.net	sare.net	wpbl.info
habeas.com	sbl.dnsbl-sh.carnet.hr	zen.dnsbl-sh.carnet.hr
hexamail.com		

Table 2: *Domains that were removed from analysis*

For each domain name that has exhibited parking behavior, we can generate a course-grained time series of the behavior to categorize what occurred. Table 3 demonstrates some sample behavioral groupings. P indicates a day where the only rdata was in private IP address space, G indicates a day where the only rdata was in globally routeable IP address space, and X indicates a day where both address types were observed, indicating a day a change between parking and active occurred.

Analysis of the domains found to exhibit parking mostly included simple text

January:	1-8	9-16	17-24	25-31
Activation on Jan 19	PPPPPPPP	PPPPPPPP	PPXGGGGG	GGGGGGG
Deactivation on Jan 19	GGGGGGGG	GGGGGGGG	GGXPPPPP	PPPPPPP
com.alextringham	GGGGGGGG	GGGGGGGG	GGGGXPPX	PXPXPPP
cn.proxyie	GGXXXXXX	GGGXGXGG	GGPGGXGX	XGGGXGX
net.homeip.bnlv	GGGGGPGG	GGGGGGGG	PGPPGGGG	GGGGPGG

Table 3: *Example parking behavior patterns, January 2014. G := only globally routable IPs observed for a domain on a given day. P := only privately reserved IPs observed. X := both observed on same day.*

matching on lists of malicious domains. While we have expressed our doubts about the soundness of evaluating an approach by comparing it to black lists [11], we have mitigated this analysis error by including as many lists as possible and limiting our assumptions of the information provided by this comparison.

Analysis of routable IP addresses includes geolocation and ASN attribution information. Geolocation is derived from the public MaxMind GeoLite2 free geolocation data from January 28, 2014 [12]. ASN attribution is derived from our publicly available IP-to-ASN mapping published for January 31, 2014,[1] itself derived from the RouteViews [13] and RIPE NCC RIS [14] data. The baseline mapping of ASNs across all IP space uses our open-source SiLK [15] tools for prefix maps and IP sets [16].

3 Results

We applied our method to all unique domains observed in our passive DNS data source for the month of January 2014. This data set contains 610 million total unique domains. After applying our method described above, 21,328 unique domains exhibit parking, or 0.0035% of the total unique domains. This number includes domains that should not publicly resolve, such as `.local`, but which did in fact have both private and public DNS answers during the period of observation.

An additional 34 domains were found to appear to exhibit parking behavior, however all 34 domains were extremely popular domains listed in the Alexa top 100 at the time [17]. We did not count these popular domains in the 21,328 that we considered to exhibit parking behavior.

In order for some assessment of known maliciousness, we checked these domains that exhibited IP address parking on private address space against 16 domain-based lists of malicious domains. 1,563 domains appeared on at least one such list

[1] `http://routeviews-mirror.cert.org/pmap/2014/01/20140131.bgp.pmap`

TLD	Count	% of Parking	% of All Domains
com	8594	40.2831%	65.7351%
net	2651	12.4262%	20.7651%
org	1045	4.89828%	1.9186%
br	842	3.94675%	0.2514%
edu	662	3.10303%	0.1268%
tw	660	3.09365%	0.0430%
ru	463	2.17024%	0.5156%
cn	441	2.06712%	0.0931%
biz	336	1.57495%	0.2265%
cc	282	1.32183%	0.2541%

Table 4: *Top 10 TLDs by number of domains exhibiting IP-address parking on private address space*

between January 1 and February 28, 2014. We allowed some additional time beyond when the domains exhibited parking in order to allow a better chance the domain would be discovered by a list, as there is some expected lag time for detection.

In order to assess some features of the network connectivity and domain structure, the 21,328 domains can be broken down by top-level domain (TLD) and whether the domain is hosted by a known dynamic DNS provider. Table 4 details the breakdown of the parking domains by TLD. We compared the 21,328 domains to a list of 71 known dynamic DNS providers as well: 353 domains were hosted in this way. The bulk hosted on two providers: 111 on dyndns.org and 191 on some name affiliated with no-ip. These are the two biggest dynamic DNS providers generally.

We can also characterize the IP addresses used to host the domains while they were routable. 41,170 unique public IP addresses were used as the routable IP addresses for some domain that exhibited parking (on private IP addresses). Each IP address had an average of 1.38 domains pointing to it, though there is clearly a heavily skewed distribution, as displayed in Table 5. We can also characterize these IP addresses by their geographic location, as best as we can determine it. The IP addresses were distributed across 164 countries, also in a long-tail distribution. Table 6 displays the 10 most common locations.

The autonomous system number (ASN) of the public IP addresses used, ASNs that announced the IP addresses were examined with the top 10 in Table 7. While the ASN counts are more evenly distributed, there is a bias of some kind towards certain ASNs. The selection of destination IP addresses is not distributed randomly across ASNs, some networks host many times the proportion of these locations

# of domains	# of IPs with X domains
$X = 1$	36765
$X \leq 10$	4169
$X \leq 50$	188
$X \leq 100$	20
$X > 100$	28

Table 5: *Distribution of domains per IP address*

Country Code	# of IP Addresses
US	17438
RU	3152
UA	2163
CN	1508
DE	1273
BR	907
CA	865
GB	809
TW	795
NL	734

Table 6: *Top 10 countries in which IP addresses of domains exhibiting parking were hosted, as geolocated on Jan 28, 2014*

than is explainable purely by chance.

4 Conclusions

The number of domains exhibiting parking on private IP addresses is quite small. And although the behavior appears to be distributed in ASNs and locations non-randomly, it does not appear to be a consistent indicator of malicious activity. The process for finding domains genuinely exhibiting parking is somewhat tedious, with a fair amount of manual review and whitelisting of domains for non-location uses that confuse the results. The process also requires a relatively long observation window, as the observation must allow enough time for the domain to change `rdata`. These two features impose a relatively high cost on finding parking domains, while there are not clear benefits to discovering them. The domains do not have a clear malicious intent, there are not many of them, and the domains are general uninteresting by our *prima facie* expert analysis. This particular kind of parking

ASN	Count	% of parking IPs	% of Internet assigned to ASN
AS6079	1574	3.82317%	0.02171%
Unknown	881	2.13991%	37.63242%
AS6517	834	2.02575%	0.00833%
AS22773	799	1.94073%	0.27731%
AS5739	732	1.77799%	0.00305%
AS8075	629	1.52781%	0.03512%
AS4134	601	1.45980%	2.52874%
AS15003	585	1.42094%	0.04291%
AS3462	525	1.27520%	0.28541%
AS46606	519	1.26063%	0.01507%

Table 7: *Top 10 ASNs announcing routable IP addresses used by domains that exhibit parking. ASN mappings are as of January 15, 2014.*

behavior does not appear to be useful to detect. The malicious behavior detected in this way would very likely be easier to detect by existing methods.

It is possible that the domains exhibiting this kind of parking are actually malicious, but simply are not found by any other method that would have them end up on the black lists we compare against. As lists of malicious behavior are mostly idiosyncratic [11], this is not entirely unlikely. We have made the complete list of domains available[2] in case another analysis can determine if they are, in fact, interesting. If so, we would welcome being proven wrong about their uninterestingness.

Acknowledgment

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

[2] `TODO--putURLhere`

References

[1] S. Alrwais, K. Yuan, E. Alowaisheq, Z. Li, and X. Wang, "Understanding the dark side of domain parking," in *23rd USENIX Security Symposium (USENIX Security 14)*. San Diego, CA: USENIX Association, Aug 2014. [Online]. Available: https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/alrwais

[2] J. Szurdi, B. Kocso, G. Cseh, J. Spring, M. Felegyhazi, and C. Kanich, "The long "taile" of typosquatting domain names," in *23rd USENIX Security Symposium (USENIX Security 14)*. San Diego, CA: USENIX Association, Aug 2014. [Online]. Available: https://www.usenix.org/conference/usenixsecurity14/technical-sessions/presentation/szurdi

[3] S. Yadav, A. K. K. Reddy, A. Reddy, and S. Ranjan, "Detecting algorithmically generated malicious domain names," in *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 2010, pp. 48–61.

[4] M. Knysz, X. Hu, and K. G. Shin, "Charlatans' web: Analysis and application of global IP-usage patterns of fast-flux botnets," *University of Michigan Ann Arbor*, pp. 1–22, 2011.

[5] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, and E. Lear, "Address Allocation for Private Internets," RFC 1918 (Best Current Practice), Tech. Rep. RFC 1918, Feb. 1996.

[6] SURBL, "Implementation guidelines," Dec 9, 2011, [Accessed: Aug 1, 2014]. [Online]. Available: http://www.surbl.org/guidelines

[7] FarSight Security, Inc., "nmsgtool," Sep 25, 2013, [Accessed: Aug 12, 2014]. [Online]. Available: https://archive.farsightsecurity.com/nmsgtool/

[8] P. Mockapetris, "Domain names - implementation and specification," RFC 1035 (Standard), Tech. Rep. RFC 1035, Nov. 1987, updated by RFCs 1101, 1183, 1348, 1876, 1982, 1995, 1996, 2065, 2136, 2181, 2137, 2308, 2535, 2845, 3425, 3658, 4033, 4034, 4035, 4343, 5936, 5966, 6604.

[9] J. Reynolds and J. Postel, "Assigned Numbers," RFC 1700 (Historic), Tech. Rep. RFC 1700, Oct. 1994, obsoleted by RFC 3232.

[10] S. Cheshire, B. Aboba, and E. Guttman, "Dynamic Configuration of IPv4 Link-Local Addresses," RFC 3927 (Proposed Standard), Tech. Rep. RFC 3927, May 2005.

[11] L. B. Metcalf and J. M. Spring, "Everything you wanted to know about blacklists but were afraid to ask," Software Engineering Institute, CERT Coordination Center, Pittsburgh, PA, Tech. Rep. CERTCC-2013-39, 2013.

[12] MaxMind, "Geolite2 free downloadable databases," Jan 28, 2014. [Online]. Available: http://dev.maxmind.com/geoip/geoip2/geolite2/

[13] Route-Views, "University of oregon route views project," http://www.routeviews.org, January 3, 2012. [Online]. Available: http://www.routeviews.org

[14] RIPE Network Coordination Center, "Routing information service (RIS)," http://www.ripe.net/data-tools/stats/ris/routing-information-service, January 3, 2012. [Online]. Available: http://www.ripe.net/data-tools/stats/ris/routing-information-service

[15] CERT/NetSA at Carnegie Mellon University, "SiLK (System for Internet-Level Knowledge)," [Accessed: Feb 4, 2014]. [Online]. Available: http://tools.netsa.cert.org/silk

[16] M. Thomas, L. Metcalf, J. Spring, P. Krystosek, and K. Prevost, "Silk: A tool suite for unsampled network flow analysis at scale," in *IEEE BigData Congress*. Anchorage, AK: IEEE, July 2014. [Online]. Available: http://resources.sei.cmu.edu/asset_files/ConferencePaper/2014_021_001_298841.pdf

[17] Alexa, "Alexa Internet, inc. – top sites," http://www.alexa.com/topsites, January 13, 2013.